

1. 通过 CM 安装 CDH

- 1.1. 为什么许多企业使用CM（Cloudera Manager）呢
- 1.2. Cloudera Manager 在企业的作用
- 1.3. 架构原理
- 1.4. 官网介绍和版本选择
- 1.5. CM对系统的要求
- 1.6. 软件下载

2. 准备集群

- 2.1. 集群规划
- 2.2. 准备四台服务器
 - 2.2.1. 修改主机名称
 - 2.2.2. 配置网络环境
 - 2.2.3. 禁用 IPV6
 - 2.2.4. 修改主机映射
 - 2.2.5. 关闭 NetworkManager
 - 2.2.6. 关闭防火墙
 - 2.2.7. 关闭 SELinux
 - 2.2.8. 设置用户最大能打开文件数目、进程数和内存
 - 2.2.9. 修改 Linux 内核参数
 - 2.2.10. 关闭透明大页面
 - 2.2.11. 安装 JDK
 - 2.2.12. 安装 MySQL
 - 2.2.13. 放置 MySQL 的驱动jar
 - 2.2.14. 执行克隆操作
- 2.3. 克隆之后必要的准备操作
 - 2.3.1. 克隆出来3台服务器
 - 2.3.2. 配置SSH免密登录

3. 安装 CM

- 3.1. 先创建安装目录
- 3.2. 复制 MySQL 驱动
- 3.3. 配置 MySQL 元数据库
- 3.4. 下载 CM 对应的安装包，解压安装
- 3.5. CDH内部用户
- 3.6. 主节点部署离线 parcel 源
- 3.7. CM 启动
 - 3.7.1. 主节点启动
 - 3.7.2. 代理节点启动
 - 3.7.3. 登录

4. 安装 CDH

5. 异常总结

1. 通过 CM 安装 CDH

1.1. 为什么许多企业使用CM（Cloudera Manager）呢

Cloudera Manager的设计目的就是针对企业，为了方便企业数据中心的管理简单和直观，在一定程度上降低了公司的成本

- 1、人员成本：减少了搭建集群的人员和维护人员；
- 2、时间成本：在一定程度上较Apache版本减少搭建时间，小白式安装,维护时间，任务运行时间，
- 3、提高了公司的资源使用，设置了资源池，有利于任务的高效处理，解决了大数据技术栈中各组件间的兼容性。同时Cloudera Manager提供了一系列的报告和诊断工具，有利于集群性能优化，提供了中央控制器对集群配置统一处理修改。

在功能上 Cloudera Manager 宕机也不会影响到其他组件的任务运行，配置存放于SQL数据库，避免了运维人员误操作导致的集群运行失败，主要得益于Cloudera Manager的架构设计。

在组件配置中类window操作，不需要写命令，看到即得到，同时他也提供了对应的API供开发者使用。

对于权限也分完全管理员和一般管理员，提高了集群的安全性，当集群出现警告会第一时间通过邮件通知，有效的降低了集群宕机的风险。

1.2. Cloudera Manager 在企业的作用

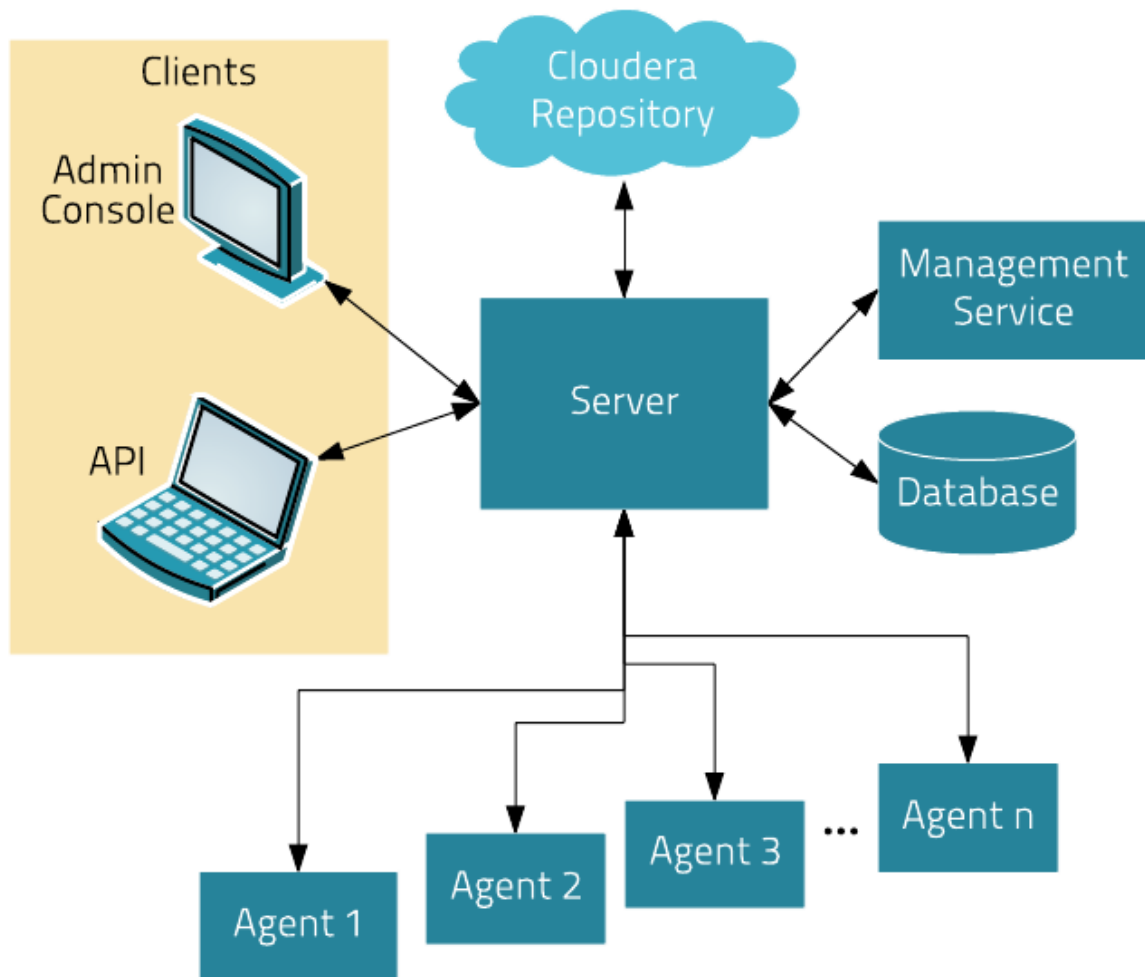
企业需要的就是这种能高效处理，把更多时间投入到开发的工具上，所以许多企业都会选择基于Cloudera Manager 监控的 CDH 版本的集群：

- 1、是有效的监控集群的健康状态
- 2、是有效的解决了hadoop生态圈中各组件及版本的兼容性。
- 3、为什么给大家推荐 Cloudera Manager，从多方面考虑：

- 1、让小白了解到目前企业使用的技术，有针对性的去学习，在学习的道路少走弯路，能早进入公司，同时公司招聘也是招聘能快速完成任务的人，不用再去学习从而降低了公司的成本。
- 2、让想从其他行业进军大数据的人员有一定的认识，方便快速进入大数据行业。
- 3、对于没有接触过Cloudera Manager的大数据人员，让其清楚并学会，如何尽可能多的时间投入开发中，以最小的成本达到效果。
- 4、让大咖了解到部分自己在实际生产环境中没有接触到的东西，认识更加全面。

1.3. 架构原理

Cloudera Manager的核心是管理服务器，该服务器承载管理控制台的Web服务器和应用程序逻辑，并负责安装软件，配置，启动和停止服务，以及管理上的服务运行群集。



Cloudera Manager Server由以下几个部分组成：

Agent: 安装在每台主机上。该代理负责启动和停止的过程，拆包配置，触发装置和监控主机。

Management Service: 由一组执行各种监控，警报和报告功能角色的服务。

Database: 存储配置和监视信息。通常情况下，多个逻辑数据库在一个或多个数据库服务器上运行。例如，Cloudera的管理服务器和监控角色使用不同的逻辑数据库。

Cloudera Repository: 软件由Cloudera 管理分布存储库。

Clients: 是用于与服务器进行交互的接口：

Admin Console : 基于web的用户界面与管理员管理集群和Cloudera管理。

API : 与开发人员创建自定义的Cloudera Manager应用程序的API。

1.4. 官网介绍和版本选择

官网: <https://www.cloudera.com/downloads/manager/5-16-2.html>

我们选择不新不旧的 CDH 稳定版本: cdh-5.16.2

1.5. CM对系统的要求

通过查阅官网，可以得知 CM 对系统的一些要求：

支持的操作系统
包括jdk的版本
包括浏览器
包括数据库
包括CDH
包括资源要求
包括网络和安全要求

和操作系统兼容: https://docs.cloudera.com/documentation/enterprise/release-notes/topics/rn_consolidated_pcm.html#c516_supported_os

和JDK兼容: https://docs.cloudera.com/documentation/enterprise/release-notes/topics/rn_consolidated_pcm.html#pcm_jdk

和MySQL兼容: https://docs.cloudera.com/documentation/enterprise/release-notes/topics/rn_consolidated_pcm.html#cdh_cm_supported_db

查看后, 总结我们准备的:

操作系统: CentOS-7.7
JDK: JDK 1.8.0_144
DataBase: MySQL-5.7
server节点: 内存至少8G
agent节点: 内存2G

软件环境

编号	软件	版本
1	操作系统	CentOS 7.7
2	JDK	jdk-8u151-linux-x64
3	Clouder Manager	5.16.2
4	CDH	5.16.2
5	数据库	MySQL-5.7
6	JDBC	mysql-connector-java-5.1.43.jar

1.6. 软件下载

JDK下载:

<https://www.oracle.com/java/technologies/javase/javase8-archive-downloads.html>

CM下载:

http://archive.cloudera.com/cm5/cm/5/cloudera-manager-centos7-cm5.16.2_x86_64.tar.gz

Index of cm5/cm/5/

Name	Last Modified	Size
Parent Directory		
cloudera-manager-centos7-cm5.10.0_x86_64.tar.gz	2018-02-21 13:20	615.94MB
cloudera-manager-centos7-cm5.10.1_x86_64.tar.gz	2018-02-21 13:20	617.43MB
cloudera-manager-centos7-cm5.10.2_x86_64.tar.gz	2018-02-21 13:20	607.17MB
cloudera-manager-centos7-cm5.11.0_x86_64.tar.gz	2018-02-21 13:20	662.24MB
cloudera-manager-centos7-cm5.11.1_x86_64.tar.gz	2018-02-21 13:20	713.57MB
cloudera-manager-centos7-cm5.11.2_x86_64.tar.gz	2018-02-21 13:21	702.66MB
cloudera-manager-centos7-cm5.12.0_x86_64.tar.gz	2018-02-21 13:21	720.49MB
cloudera-manager-centos7-cm5.12.1_x86_64.tar.gz	2018-02-21 13:21	708.30MB
cloudera-manager-centos7-cm5.12.2_x86_64.tar.gz	2018-02-21 13:21	722.32MB
cloudera-manager-centos7-cm5.13.0_x86_64.tar.gz	2018-02-21 13:21	718.79MB
cloudera-manager-centos7-cm5.13.1_x86_64.tar.gz	2018-02-21 13:21	718.81MB
cloudera-manager-centos7-cm5.13.2_x86_64.tar.gz	2018-02-21 13:21	715.87MB
cloudera-manager-centos7-cm5.13.3_x86_64.tar.gz	2018-04-04 10:53	715.97MB
cloudera-manager-centos7-cm5.14.0_x86_64.tar.gz	2018-02-21 13:21	749.38MB
cloudera-manager-centos7-cm5.14.1_x86_64.tar.gz	2018-02-21 13:21	749.39MB
cloudera-manager-centos7-cm5.14.2_x86_64.tar.gz	2018-04-09 13:35	793.90MB
cloudera-manager-centos7-cm5.14.3_x86_64.tar.gz	2018-04-18 14:40	793.92MB
cloudera-manager-centos7-cm5.14.4_x86_64.tar.gz	2018-07-12 18:44	797.50MB
cloudera-manager-centos7-cm5.15.0_x86_64.tar.gz	2018-06-14 09:20	798.53MB
cloudera-manager-centos7-cm5.15.1_x86_64.tar.gz	2018-08-22 14:48	800.03MB
cloudera-manager-centos7-cm5.15.2_x86_64.tar.gz	2018-12-13 16:21	801.46MB
cloudera-manager-centos7-cm5.16.1_x86_64.tar.gz	2018-11-27 14:37	802.54MB
cloudera-manager-centos7-cm5.16.2_x86_64.tar.gz	2019-06-18 13:09	799.13MB
cloudera-manager-centos7-cm5.4.8_x86_64.tar.gz	2018-02-21 13:21	660.15MB
cloudera-manager-centos7-cm5.5.0_x86_64.tar.gz	2018-02-21 13:22	533.87MB
cloudera-manager-centos7-cm5.5.1_x86_64.tar.gz	2018-02-21 13:22	493.62MB

CDH下载:

<http://archive.cloudera.com/cdh5/parcels/5.16.2/CDH-5.16.2-1.cdh5.16.2.p0.8-el7.parcel>
<http://archive.cloudera.com/cdh5/parcels/5.16.2/CDH-5.16.2-1.cdh5.16.2.p0.8-el7.parcel.sha1>
<http://archive.cloudera.com/cdh5/parcels/5.16.2/manifest.json>

Index of cdh5/parcels/5.16.2/

Name	Last Modified	Size
Parent Directory		
CDH-5.16.2-1.cdh5.16.2.p0.8-el5.parcel	2019-06-18 13:18	1.57GB
CDH-5.16.2-1.cdh5.16.2.p0.8-el5.parcel.sha1	2019-06-18 13:17	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-el6.parcel	2019-06-18 13:18	1.74GB
CDH-5.16.2-1.cdh5.16.2.p0.8-el6.parcel.sha1	2019-06-18 13:18	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-el7.parcel	2019-06-18 13:17	1.99GB
CDH-5.16.2-1.cdh5.16.2.p0.8-el7.parcel.sha1	2019-06-18 13:18	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-jessie.parcel	2019-06-18 13:17	1.84GB
CDH-5.16.2-1.cdh5.16.2.p0.8-jessie.parcel.sha1	2019-06-18 13:18	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-sles11.parcel	2019-06-18 13:18	1.58GB
CDH-5.16.2-1.cdh5.16.2.p0.8-sles11.parcel.sha1	2019-06-18 13:18	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-sles12.parcel	2019-06-18 13:18	1.73GB
CDH-5.16.2-1.cdh5.16.2.p0.8-sles12.parcel.sha1	2019-06-18 13:18	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-trusty.parcel	2019-06-18 13:18	1.84GB
CDH-5.16.2-1.cdh5.16.2.p0.8-trusty.parcel.sha1	2019-06-18 13:18	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-wheezy.parcel	2019-06-18 13:17	1.70GB
CDH-5.16.2-1.cdh5.16.2.p0.8-wheezy.parcel.sha1	2019-06-18 13:18	41B
CDH-5.16.2-1.cdh5.16.2.p0.8-xenial.parcel	2019-06-18 13:18	1.85GB
CDH-5.16.2-1.cdh5.16.2.p0.8-xenial.parcel.sha1	2019-06-18 13:18	41B
manifest.json	2019-06-18 13:18	65.24KB

MySQL驱动下载:

<http://central.maven.org/maven2/mysql/mysql-connector-java/5.1.43/mysql-connector-java-5.1.43.jar>

2. 准备集群

2.1. 集群规划

我们准备4台虚拟机，我们需要准备一台内存至少为16G的主机，然后在主机上虚拟4台虚拟机、虚拟机的主机名和内存分别是：

NO	机器名称	内网IP	配置	用途
1	bigdata02	192.168.123.102	8C/32Gb/1TB	Master, CM, MySQL
2	bigdata03	192.168.123.103	8C/32Gb/1TB	Agent
3	bigdata04	192.168.123.104	8C/32Gb/1TB	Agent
4	bigdata05	192.168.123.105	8C/32Gb/1TB	Agent

说明：CDH推荐的 worker 的数量最好是3个及以上。所以我们准备3个。

可以参考CDH官网推荐：https://docs.cloudera.com/documentation/enterprise/latest/topics/cm_ig_host_allocations.html#concept_f43_j4y_dw_section_icy_mgj_ndb

可以参考CSDN文章博客理解：<https://blog.csdn.net/chenguangchun1993/article/details/79164857>

两篇文章，第一篇是CDH官网推荐的角色分配、第二篇是CSDN上的博文

2.2. 准备四台服务器

2.2.1. 修改主机名称

2.2.2. 配置网络环境

2.2.3. 禁用 IPV6

2.2.4. 修改主机映射

2.2.5. 关闭 NetworkManager

执行如下命令：

```
systemctl stop NetworkManager
systemctl disable NetworkManager
```

2.2.6. 关闭防火墙

2.2.7. 关闭 SELinux

2.2.8. 设置用户最大能打开文件数目、进程数和内存

2.2.9. 修改 Linux 内核参数

设置 swappiness, 控制换出运行时内存的相对权重, Cloudera 建议将 swappiness 设置为 10:

```
# 查看swappiness
cat /proc/sys/vm/swappiness

# 永久性修改, 执行下面两条命令
sysctl -w vm.swappiness=10
echo vm.swappiness = 10 >> /etc/sysctl.conf
```

2.2.10. 关闭透明大页面

2.2.11. 安装 JDK

2.2.12. 安装 MySQL

2.2.13. 放置 MySQL 的驱动jar

2.2.14. 执行克隆操作

2.3. 克隆之后必要的准备操作

2.3.1. 克隆出来3台服务器

2.3.2. 配置SSH免密登录

3. 安装 CM

3.1. 先创建安装目录

创建各种所需目录:

```
mkdir -p /opt/cloudera/parcels
mkdir -p /opt/cloudera/parcel-repo
mkdir -p /opt/cloudera/parcel-cache
mkdir -p /opt/cloudera/csd
mkdir -p /opt/cloudera-manager
mkdir -p /usr/share/java/
```

3.2. 复制 MySQL 驱动

复制 MySQL JDBC 驱动 Jar 包到 /usr/share/java/ 目录

```
cd /usr/share/java/
```

```
cp ~/soft/mysql-connector-java-5.1.45-bin.jar /usr/share/java/
```

并且需要改名：

```
mv /usr/share/java/mysql-connector-java-5.1.45-bin.jar /usr/share/java/mysql-connector-java.jar
```

所有节点都需要！

```
scp -r /usr/share/java/mysql-connector-java.jar bigdata12_cdh02:/usr/share/java/  
scp -r /usr/share/java/mysql-connector-java.jar bigdata13_cdh03:/usr/share/java/  
scp -r /usr/share/java/mysql-connector-java.jar bigdata14_cdh04:/usr/share/java/
```

3.3. 配置 MySQL 元数据库

创建ClouderaManager需要用到的几个数据库

```
mysql -uroot -pQwer_1234
```

```
create database hive DEFAULT CHARSET utf8 COLLATE utf8_general_ci;  
create database amon DEFAULT CHARSET utf8 COLLATE utf8_general_ci;  
create database rm DEFAULT CHARSET utf8 COLLATE utf8_general_ci;  
create database hue DEFAULT CHARSET utf8 COLLATE utf8_general_ci;  
create database monitor DEFAULT CHARSET utf8 COLLATE utf8_general_ci;  
create database oozie DEFAULT CHARSET utf8 COLLATE utf8_general_ci;  
create database cmf DEFAULT CHARSET utf8 COLLATE utf8_general_ci;  
create database ambari DEFAULT CHARSET utf8 COLLATE utf8_general_ci;
```

3.4. 下载 CM 对应的安装包，解压安装

到官网下载相对应的安装包：http://archive.cloudera.com/cm5/cm/5/cloudera-manager-el6-cm5.3.6_x86_64.tar.gz

把 cloudera-manager-centos7-cm5.16.2_x86_64.tar.gz 解压到 /opt/cloudera-manager（所有节点）

```
tar -zxvf ~/soft/cloudera-manager-centos7-cm5.16.1_x86_64.tar.gz -C  
/opt/cloudera-manager
```

修改所有主机 /opt/cloudera-manager/cm-5.16.1/etc/cloudera-scm-agent 下的 config.ini，只需要修改这一行

```
vim /opt/cloudera-manager/cm-5.16.1/etc/cloudera-scm-agent/config.ini
```

把 server_host 做更改：设置成你想设置的 CM 的主节点。则其他节点都是代理节点。

```
server_host=bigdata12
```


然后再来修改CM要使用的元数据库的相关信息。

```
vim /opt/cloudera-manager/cm-5.16.1/etc/cloudera-scm-server/db.properties
```

```
com.cloudera.cmf.db.type=mysql
com.cloudera.cmf.db.host=bigdata12
com.cloudera.cmf.db.name=cmf
com.cloudera.cmf.db.user=root
com.cloudera.cmf.db.password=Qwer_1234
com.cloudera.cmf.db.setupType=EXTERNAL
```

所有的节点都需要进行同步一下：

```
scp -r /opt/cloudera-manager/cm-5.16.1 bigdata12:/opt/cloudera-manager/
scp -r /opt/cloudera-manager/cm-5.16.1 bigdata13:/opt/cloudera-manager/
scp -r /opt/cloudera-manager/cm-5.16.1 bigdata14:/opt/cloudera-manager/
```

3.5. CDH内部用户

所有节点都需要做这个操作！

创建用户：

```
useradd --system --home=/opt/cloudera-manager/cm-5.16.1/run/cloudera-scm-server/
--no-create-home --shell=/bin/false --comment "Cloudera SCM User" cloudera-scm
```

修改权限：

```
chown -R cloudera-scm:cloudera-scm /opt/cloudera-manager
chown -R cloudera-scm:cloudera-scm /opt/cloudera
```

3.6. 主节点部署离线 parcel 源

```
mkdir -p /opt/cloudera/parcel-repo
```

把 `CDH-5.16.1-1.cdh5.16.1.p0.3-e17.parcel` 和 `CDH-5.16.1-1.cdh5.16.1.p0.3-e17.parcel.sha1` 和 `manifest.json` 放到 `/opt/cloudera/parcel-repo` 文件夹中。

记得 `CDH-5.16.1-1.cdh5.16.1.p0.3-e17.parcel.sha1` 要改名，把后面的 1 去掉！

```
mv CDH-5.16.1-1.cdh5.16.1.p0.3-e17.parcel.sha1 CDH-5.16.1-1.cdh5.16.1.p0.3-e17.parcel.sha
```

修改权限：

```
chown -R cloudera-scm:cloudera-scm /opt/cloudera/
```

```
cd /opt/cloudera/parcel-repo
11
```

3.7. CM 启动

3.7.1. 主节点启动

```
/opt/cloudera-manager/cm-5.16.1/etc/init.d/cloudera-scm-server start
```

看日志：

```
tail -f /opt/cloudera-manager/cm-5.16.1/log/cloudera-scm-server/cloudera-scm-server.log
```

3.7.2. 代理节点启动

```
/opt/cloudera-manager/cm-5.16.1/etc/init.d/cloudera-scm-agent start
```

看日志：

```
cd /opt/cloudera-manager/cm-5.16.1/log/cloudera-scm-agent
```

3.7.3. 登录

登录页面：<http://bigdata12:7180/cm/login>

用户：admin

密码：admin

cloudera manager支持 · 1.5

欢迎使用 Cloudera Manager。您想要部署哪一个版本？

升级到 Cloudera Enterprise Data Hub Edition 将提供可以帮助您在关键任务环境下管理和监控 Hadoop 集群的重要功能。

	Cloudera Express	Cloudera Enterprise 数据集线器试用版	Cloudera Enterprise
许可证	✓ 免费	60 天 在试用期之后，该产品将继续作为 Cloudera Express 运行。您的群集和数据将 会保持不受影响。	年度订阅 下载许可证 Cloudera Enterprise 在三个版本中可用： <ul style="list-style-type: none">Basic EditionFlex EditionData Hub Edition
节点限制	无限制	无限制	无限制
CDH	✓	✓	✓
Cloudera Manager 核心功能	✓	✓	✓
Cloudera Manager 高级功能		✓	✓
Cloudera Navigator		✓	✓
Cloudera 支持			✓

如需获得 Cloudera Express 和 Cloudera Enterprise 中提供的功能的完整列表，[单击此处](#)。

[继续](#)

4. 安装 CDH

具体看操作!

5. 异常总结

- 1、主机名称，千万不能带 "_"，否则你会哭的。
- 2、HUE 元数据库测试失败：请安装 mysql 的兼容库

```
rpm -ivh mysql-community-libs-compat-5.7.30-1.el7.x86_64.rpm
```